2

Arguments Concerning Scientific Realism

> The rigour of science requires that we distinguish well the undraped figure of nature itself from the gay-coloured vesture with which we clothe it at our pleasure.
>
> Heinrich Hertz, quoted by Ludwig Boltzmann, letter to *Nature*, 28 February 1895

In our century, the first dominant philosophy of science was developed as part of logical positivism. Even today, such an expression as 'the received view of theories' refers to the views developed by the logical positivists, although their heyday preceded the Second World War.

In this chapter I shall examine, and criticize, the main arguments that have been offered for scientific realism. These arguments occurred frequently as part of a critique of logical positivism. But it is surely fair to discuss them in isolation, for even if scientific realism is most easily understood as a reaction against positivism, it should be able to stand alone. The alternative view which I advocate —for lack of a traditional name I shall call it *constructive empiricism* —is equally at odds with positivist doctrine.

§1. *Scientific Realism and Constructive Empiricism*

In philosophy of science, the term 'scientific realism' denotes a precise position on the question of how a scientific theory is to be understood, and what scientific activity really is. I shall attempt to define this position, and to canvass its possible alternatives. Then I shall indicate, roughly and briefly, the specific alternative which I shall advocate and develop in later chapters.

§1.1 *Statement of Scientific Realism*

What exactly is scientific realism? A naïve statement of the position would be this: the picture which science gives us of the world is a

true one, faithful in its details, and the entities postulated in science really exist: the advances of science are discoveries, not inventions. That statement is too naïve; it attributes to the scientific realist the belief that today's theories are correct. It would mean that the philosophical position of an earlier scientific realist such as C. S. Peirce had been refuted by empirical findings. I do not suppose that scientific realists wish to be committed, as such, even to the claim that science will arrive in due time at theories true in all respects—for the growth of science might be an endless self-correction; or worse, Armageddon might occur too soon.

But the naïve statement has the right flavour. It answers two main questions: it characterizes a scientific theory as a story about what there really is, and scientific activity as an enterprise of discovery, as opposed to invention. The two questions of what a scientific theory is, and what a scientific theory does, must be answered by any philosophy of science. The task we have at this point is to find a statement of scientific realism that shares these features with the naïve statement, but does not saddle the realists with unacceptably strong consequences. It is especially important to make the statement as weak as possible if we wish to argue against it, so as not to charge at windmills.

As clues I shall cite some passages most of which will also be examined below in the contexts of the authors' arguments. A statement of Wilfrid Sellars is this:

> to have good reason for holding a theory is *ipso facto* to have good reason for holding that the entities postulated by the theory exist.[7]

This addresses a question of epistemology, but also throws some indirect light on what it is, in Sellars's opinion, to hold a theory. Brian Ellis, who calls himself a scientific entity realist rather than a scientific realist, appears to agree with that statement of Sellars, but gives the following formulation of a stronger view:

> I understand scientific realism to be the view that the theoretical statements of science are, or purport to be, true generalized descriptions of reality.[1]

This formulation has two advantages: It focuses on the understanding of the theories without reference to reasons for belief, and it avoids the suggestion that to be a realist you must believe current scientific theories to be true. But it gains the latter advantage by use of the word 'purport', which may generate its own puzzles.

Hilary Putnam, in a passage which I shall cite again in Section 7, gives a formulation which he says he learned from Michael Dummett:

> A realist (with respect to a given theory or discourse) holds that (1) the sentences of that theory are true or false; and (2) that what makes them true or false is something external—that is to say, it is not (in general) our sense data, actual or potential, or the structure of our minds, or our language, etc.[29]

He follows this soon afterwards with a further formulation which he credits to Richard Boyd:

> That terms in mature scientific theories typically refer (this formulation is due to Richard Boyd), that the theories accepted in a mature science are typically approximately true, that the same term can refer to the same thing even when it occurs in different theories—these statements are viewed by the scientific realist ... as part of any adequate scientific description of science and its relations to its objects.[33]

None of these were intended as definitions. But they show I think that truth must play an important role in the formulation of the basic realist position. They also show that the formulation must incorporate an answer to the question what it is to *accept* or *hold* a theory. I shall now propose such a formulation, which seems to me to make sense of the above remarks, and also renders intelligible the reasoning by realists which I shall examine below—without burdening them with more than the minimum required for this.

Science aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true. This is the correct statement of scientific realism.

Let me defend this formulation by showing that it is quite minimal, and can be agreed to by anyone who considers himself a scientific realist. The naïve statement said that science tells a true story; the correct statement says only that it is the aim of science to do so. The aim of science is of course not to be identified with individual scientists' motives. The aim of the game of chess is to checkmate your opponent; but the motive for playing may be fame, gold, and glory. What the aim is determines what counts as success in the enterprise as such; and this aim may be pursued for any number of reasons. Also, in calling something *the* aim, I do not deny that there are other subsidiary aims which may or may not be means to that end: everyone will readily agree that simplicity, informativeness, predictive power, explanation are (also) virtues. Perhaps my formulation can even be accepted by any philosopher who considers the most important aim of science to be something which only *requires* the finding of true theories—given that I wish to give the weakest formulation of the doctrine that is generally acceptable.

I have added 'literally' to rule out as realist such positions as imply that science is true if 'properly understood' but literally false or meaningless. For that would be consistent with conventionalism, logical positivism, and instrumentalism. I will say more about this below; and also in Section 7 where I shall consider Dummett's views further.

The second part of the statement touches on epistemology. But it only equates acceptance of a theory with belief in its truth.[2] It does not imply that anyone is ever rationally warranted in forming such a belief. We have to make room for the epistemological position, today the subject of considerable debate, that a rational person never assigns personal probability 1 to any proposition except a tautology. It would, I think, be rare for a scientific realist to take this stand in epistemology, but it is certainly possible.[3]

To understand qualified acceptance we must first understand acceptance *tout court*. If acceptance of a theory involves the belief that it is true, then tentative acceptance involves the tentative adoption of the belief that it is true. If belief comes in degrees, so does acceptance, and we may then speak of a degree of acceptance involving a certain degree of belief that the theory is true. This must of course be distinguished from belief that the theory is approximately true, which seems to mean belief that some member of a class centring on the mentioned theory is (exactly) true. In this way the proposed formulation of realism can be used regardless of one's epistemological persuasion.

§1.2 *Alternatives to Realism*

Scientific realism is the position that scientific theory construction aims to give us a literally true story of what the world is like, and that acceptance of a scientific theory involves the belief that it is true. Accordingly, anti-realism is a position according to which the aim of science can well be served without giving such a literally true story, and acceptance of a theory may properly involve something less (or other) than belief that it is true.

What does a scientist do then, according to these different positions? According to the realist, when someone proposes a theory, he

is asserting it to be true. But according to the anti-realist, the proposer does not assert the theory to be true; *he displays it*, and claims certain virtues for it. These virtues may fall short of truth: empirical adequacy, perhaps; comprehensiveness, acceptability for various purposes. This will have to be spelt out, for the details here are not determined by the denial of realism. For now we must concentrate on the key notions that allow the generic division.

The idea of a literally true account has two aspects: the language is to be literally construed; and so construed, the account is true. This divides the anti-realists into two sorts. The first sort holds that science is or aims to be true, properly (but not literally) construed. The second holds that the language of science should be literally construed, but its theories need not be true to be good. The anti-realism I shall advocate belongs to the second sort.

It is not so easy to say what is meant by a literal construal. The idea comes perhaps from theology, where fundamentalists construe the Bible literally, and liberals have a variety of allegorical, metaphorical, and analogical interpretations, which 'demythologize'. The problem of explicating 'literal construal' belongs to the philosophy of language. In Section 7 below, where I briefly examine some of Michael Dummett's views, I shall emphasize that 'literal' does not mean 'truth-valued'. The term 'literal' is well enough understood for general philosophical use, but if we try to explicate it we find ourselves in the midst of the problem of giving an adequate account of natural language. It would be bad tactics to link an inquiry into science to a commitment to some solution to that problem. The following remarks, and those in Section 7, should fix the usage of 'literal' sufficiently for present purposes.

The decision to rule out all but literal construals of the language of science, rules out those forms of anti-realism known as *positivism* and *instrumentalism*. First, on a literal construal, the apparent statements of science really are statements, *capable of* being true or false. Secondly, although a literal construal can elaborate, it cannot change logical relationships. (It is possible to elaborate, for instance, by identifying what the terms designate. The 'reduction' of the language of phenomenological thermodynamics to that of statistical mechanics is like that: bodies of gas are identified as aggregates of molecules, temperature as mean kinetic energy, and so on.) On the positivists' interpretation of science, theoretical terms have meaning only through their connection with the observable. Hence they hold

that two theories may in fact *say the same thing* although in form they contradict each other. (Perhaps the one says that all matter consists of atoms, while the other postulates instead a universal continuous medium; they will say the same thing nevertheless if they agree in their observable consequences, according to the positivists.) But two theories which contradict each other in such a way can 'really' be saying the same thing only if they are not literally construed. Most specifically, if a theory says that something exists, then a literal construal may elaborate on what that something is, but will not remove the implication of existence.

There have been many critiques of positivist interpretations of science, and there is no need to repeat them. I shall add some specific criticisms of the positivist approach in the next chapter.

§1.3 *Constructive Empiricism*

To insist on a literal construal of the language of science is to rule out the construal of a theory as a metaphor or simile, or as intelligible only after it is 'demythologized' or subjected to some other sort of 'translation' that does not preserve logical form. If the theory's statements include 'There are electrons', then the theory says that there are electrons. If in addition they include 'Electrons are not planets', then the theory says, in part, that there are entities other than planets.

But this does not settle very much. It is often not at all obvious whether a theoretical term refers to a concrete entity or a mathematical entity. Perhaps one tenable interpretation of classical physics is that there are no concrete entities which are forces—that 'there are forces such that . . .' can always be understood as a mathematical statement asserting the existence of certain functions. That is debatable.

Not every philosophical position concerning science which insists on a literal construal of the language of science is a realist position. For this insistence relates not at all to our epistemic attitudes toward theories, nor to the aim we pursue in constructing theories, but only to the correct understanding of *what a theory says*. (The fundamentalist theist, the agnostic, and the atheist presumably agree with each other (though not with liberal theologians) in their understanding of the statement that God, or gods, or angels exist.) After deciding that the language of science must be literally understood, we can still say that there is no need to believe good theories to be

true, nor to believe *ipso facto* that the entities they postulate are real.

Science aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate. This is the statement of the anti-realist position I advocate; I shall call it *constructive empiricism.*

This formulation is subject to the same qualifying remarks as that of scientific realism in Section 1.1 above. In addition it requires an explication of 'empirically adequate'. For now, I shall leave that with the preliminary explication that a theory is empirically adequate exactly if what it says about the observable things and events in this world, is true—exactly if it 'saves the phenomena'. A little more precisely: such a theory has at least one model that all the actual phenomena fit inside. I must emphasize that this refers to *all* the phenomena; these are not exhausted by those actually observed, nor even by those observed at some time, whether past, present, or future. The whole of the next chapter will be devoted to the explication of this term, which is intimately bound up with our conception of the structure of a scientific theory.

The distinction I have drawn between realism and anti-realism, in so far as it pertains to acceptance, concerns only how much belief is involved therein. Acceptance of theories (whether full, tentative, to a degree, etc.) is a phenomenon of scientific activity which clearly involves more than belief. One main reason for this is that we are never confronted with a complete theory. So if a scientist accepts a theory, he thereby involves himself in a certain sort of research programme. That programme could well be different from the one acceptance of another theory would have given him, even if those two (very incomplete) theories are equivalent to each other with respect to everything that is observable—in so far as they go.

Thus acceptance involves not only belief but a certain commitment. Even for those of us who are not working scientists, the acceptance involves a commitment to confront any future phenomena by means of the conceptual resources of this theory. It determines the terms in which we shall seek explanations. If the acceptance is at all strong, it is exhibited in the person's assumption of the role of explainer, in his willingness to answer questions *ex cathedra.* Even if you do not accept a theory, you can engage in discourse in a context in which language use is guided by that theory—but acceptance produces such contexts. There are similarities in all of this to ideological commitment. A commitment is of course not true or false: The confidence exhibited is that it will be *vindicated.*

This is a preliminary sketch of the *pragmatic* dimension of theory acceptance. Unlike the epistemic dimension, it does not figure overtly in the disagreement between realist and anti-realist. But because the amount of belief involved in acceptance is typically less according to anti-realists, they will tend to make more of the pragmatic aspects. It is as well to note here the important difference. Belief that a theory is true, or that it is empirically adequate, does not imply, and is not implied by, belief that full acceptance of the theory will be vindicated. To see this, you need only consider here a person who has quite definite beliefs about the future of the human race, or about the scientific community and the influences thereon and practical limitations we have. It might well be, for instance, that a theory which is empirically adequate will not combine easily with some other theories which we have accepted in fact, or that Armageddon will occur before we succeed. Whether belief that a theory is true, or that it is empirically adequate, can be equated with belief that acceptance of it would, under ideal research conditions, be vindicated in the long run, is another question. It seems to me an irrelevant question within philosophy of science, because an affirmative answer would not obliterate the distinction we have already established by the preceding remarks. (The question may also assume that counterfactual statements are objectively true or false, which I would deny.)

Although it seems to me that realists and anti-realists need not disagree about the pragmatic aspects of theory acceptance, I have mentioned it here because I think that typically they do. We shall find ourselves returning time and again, for example, to requests for explanation to which realists typically attach an objective validity which anti-realists cannot grant.

§2. *The Theory/Observation 'Dichotomy'*

For good reasons, logical positivism dominated the philosophy of science for thirty years. In 1960, the first volume of *Minnesota Studies in the Philosophy of Science* published Rudolf Carnap's 'The Methodological Status of Theoretical Concepts', which is, in many ways, the culmination of the positivist programme. It interprets science by relating it to an observation language (a postulated part

of natural language which is devoid of theoretical terms). Two years later this article was followed in the same series by Grover Maxwell's 'The Ontological Status of Theoretical Entities', in title and theme a direct counter to Carnap's. This is the *locus classicus* for the new realists' contention that the theory/observation distinction cannot be drawn.

I shall examine some of Maxwell's points directly, but first a general remark about the issue. Such expressions as 'theoretical entity' and 'observable–theoretical dichotomy' are, on the face of it, examples of category mistakes. Terms or concepts are theoretical (introduced or adapted for the purposes of theory construction); entities are observable or unobservable. This may seem a little point, but it separates the discussion into two issues. Can we divide our language into a theoretical and non-theoretical part? On the other hand, can we classify objects and events into observable and unobservable ones?

Maxwell answers both questions in the negative, while not distinguishing them too carefully. On the first, where he can draw on well-known supportive essays by Wilfrid Sellars and Paul Feyerabend, I am in total agreement. All our language is thoroughly theory-infected. If we could cleanse our language of theory-laden terms, beginning with the recently introduced ones like 'VHF receiver', continuing through 'mass' and 'impulse' to 'element' and so on into the prehistory of language formation, we would end up with nothing useful. The way we talk, and scientists talk, is guided by the pictures provided by previously accepted theories. This is true also, as Duhem already emphasized, of experimental reports. Hygienic reconstructions of language such as the positivists envisaged are simply not on. I shall return to this criticism of positivism in the next chapter.

But does this mean that we must be scientific realists? We surely have more tolerance of ambiguity than that. The fact that we let our language be guided by a given picture, at some point, does not show how much we believe about that picture. When we speak of the sun coming up in the morning and setting at night, we are guided by a picture now explicitly disavowed. When Milton wrote *Paradise Lost* he deliberately let the old geocentric astronomy guide his poem, although various remarks in passing clearly reveal his interest in the new astronomical discoveries and speculations of his time. These are extreme examples, but show that no immediate

conclusions can be drawn from the theory-ladenness of our language.

However, Maxwell's main arguments are directed against the observable–unobservable distinction. Let us first be clear on what this distinction was supposed to be. The term 'observable' classifies putative entities (entities which may or may not exist). A flying horse is observable—that is why we are so sure that there aren't any—and the number seventeen is not. There is supposed to be a correlate classification of human acts: an unaided act of perception, for instance, is an observation. A calculation of the mass of a particle from the deflection of its trajectory in a known force field, is not an observation of that mass.

It is also important here not to confuse *observing* (an entity, such as a thing, event, or process) and *observing that* (something or other is the case). Suppose one of the Stone Age people recently found in the Philippines is shown a tennis ball or a car crash. From his behaviour, we see that he has noticed them; for example, he picks up the ball and throws it. But he has not seen *that* it is a tennis ball, or *that* some event is a car crash, for he does not even have those concepts. He cannot get that information through perception; he would first have to learn a great deal. To say that he does not see the same things and events as we do, however, is just silly; it is a pun which trades on the ambiguity between seeing and seeing that. (The truth-conditions for our statement '*x* observes *that A*' must be such that what concepts *x* has, presumably related to the language *x* speaks if he is human, enter as a variable into the correct truth definition, in some way. To say that *x* observed the tennis ball, therefore, does not imply at all that *x* observed that it was a tennis ball; that would require some conceptual awareness of the game of tennis.)

The arguments Maxwell gives about observability are of two sorts: one directed against the possibility of drawing such distinctions, the other against the importance that could attach to distinctions that can be drawn.

The first argument is from the continuum of cases that lie between direct observation and inference:

> there is, in principle, a continuous series beginning with looking through a vacuum and containing these as members: looking through a windowpane, looking through glasses, looking through binoculars, looking through a low-power microscope, looking through a high-power microscope, etc., in the

order given. The important consequence is that, so far, we are left without criteria which would enable us to draw a non-arbitrary line between 'observation' and 'theory'.[4]

This continuous series of supposed acts of observation does not correspond directly to a continuum in what is supposed observable. For if something can be seen through a window, it can also be seen with the window raised. Similarly, the moons of Jupiter can be seen through a telescope; but they can also be seen without a telescope if you are close enough. That something is observable does not automatically imply that the conditions are right for observing it now. The principle is:

> X is observable if there are circumstances which are such that, if X is present to us under those circumstances, then we observe it.

This is not meant as a definition, but only as a rough guide to the avoidance of fallacies.

We may still be able to find a continuum in what is supposed detectable: perhaps some things can only be detected with the aid of an optical microscope, at least; perhaps some require an electron microscope, and so on. Maxwell's problem is: where shall we draw the line between what is observable and what is only detectable in some more roundabout way?

Granted that we cannot answer this question without arbitrariness, what follows? That 'observable' is a *vague predicate*. There are many puzzles about vague predicates, and many sophisms designed to show that, in the presence of vagueness, no distinction can be drawn at all. In Sextus Empiricus, we find the argument that incest is not immoral, for touching your mother's big toe with your little finger is not immoral, and all the rest differs only by degree. But predicates in natural language are almost all vague, and there is no problem in their use; only in formulating the logic that governs them.[5] A vague predicate is usable provided it has clear cases and clear counter-cases. Seeing with the unaided eye is a clear case of observation. Is Maxwell then perhaps challenging us to present a clear counter-case? Perhaps so, for he says 'I have been trying to support the thesis that any (non-logical) term is a *possible* candidate for an observation term.'

A look through a telescope at the moons of Jupiter seems to me a clear case of observation, since astronauts will no doubt be able to see them as well from close up. But the purported observation

of micro-particles in a cloud chamber seems to me a clearly different case—if our theory about what happens there is right. The theory says that if a charged particle traverses a chamber filled with saturated vapour, some atoms in the neighbourhood of its path are ionized. If this vapour is decompressed, and hence becomes supersaturated, it condenses in droplets on the ions, thus marking the path of the particle. The resulting silver-grey line is similar (physically as well as in appearance) to the vapour trail left in the sky when a jet passes. Suppose I point to such a trail and say: 'Look, there is a jet!'; might you not say: 'I see the vapour trail, but where is the jet?' Then I would answer: 'Look just a bit ahead of the trail ... there! Do you see it?' Now, in the case of the cloud chamber this response is not possible. So while the particle is detected by means of the cloud chamber, and the detection is based on observation, it is clearly not a case of the particle's being observed.

As a second argument, Maxwell directs our attention to the 'can' in 'what is observable is what can be observed.' An object might of course be temporarily unobservable—in a rather different sense: it cannot be observed in the circumstances in which it actually is at the moment, but could be observed if the circumstances were more favourable. In just the same way, I might be temporarily invulnerable or invisible. So we should concentrate on 'observable' *tout court*, or on (as he prefers to say) 'unobservable in principle'. This Maxwell explains as meaning that the relevant scientific theory *entails* that the entities cannot be observed in any circumstances. But this never happens, he says, because the different circumstances could be ones in which we have different sense organs—electron-microscope eyes, for instance.

This strikes me as a trick, a change in the subject of discussion. I have a mortar and pestle made of copper and weighing about a kilo. Should I call it breakable because a giant could break it? Should I call the Empire State Building portable? Is there no distinction between a portable and a console record player? The human organism is, from the point of view of physics, a certain kind of measuring apparatus. As such it has certain inherent limitations—which will be described in detail in the final physics and biology. It is these limitations to which the 'able' in 'observable' refers—our limitations, *qua* human beings.

As I mentioned, however, Maxwell's article also contains a different sort of argument: even if there is a feasible observable/

unobservable distinction, this distinction has no importance. The point at issue for the realist is, after all, the reality of the entities postulated in science. Suppose that these entities could be classified into observables and others; what relevance should that have to the question of their existence?

Logically, none. For the term 'observable' classifies putative entities, and has logically nothing to do with existence. But Maxwell must have more in mind when he says: 'I conclude that the drawing of the observational–theoretical line at any given point is an accident and a function of our physiological make-up, ... and, therefore, that it has no ontological significance whatever.'[6] No ontological significance if the question is only whether 'observable' and 'exists' imply each other—for they do not; but significance for the question of scientific realism?

Recall that I defined scientific realism in terms of the aim of science, and epistemic attitudes. The question is what aim scientific activity has, and how much we shall believe when we accept a scientific theory. What is the proper form of acceptance: belief that the theory, as a whole, is true; or something else? To this question, what is observable by us seems eminently relevant. Indeed, we may attempt an answer at this point: to accept a theory is (for us) to believe that it is empirically adequate—that what the theory says *about what is observable* (by us) is true.

It will be objected at once that, on this proposal, what the anti-realist decides to believe about the world will depend in part on what he believes to be his, or rather the epistemic community's, accessible range of evidence. At present, we count the human race as the epistemic community to which we belong; but this race may mutate, or that community may be increased by adding other animals (terrestrial or extra-terrestrial) through relevant ideological or moral decisions ('to count them as persons'). Hence the anti-realist would, on my proposal, have to accept conditions of the form

> If the epistemic community changes in fashion Y, then my beliefs about the world will change in manner Z.

To see this as an objection to anti-realism is to voice the requirement that our epistemic policies should give the same results independent of our beliefs about the range of evidence accessible to us. That requirement seems to me in no way rationally compelling; it could be honoured, I should think, only through a thoroughgoing scepticism or through a commitment to wholesale leaps of faith. But we cannot settle the major questions of epistemology *en passant* in philosophy of science; so I shall just conclude that it is, on the face of it, not irrational to commit oneself only to a search for theories that are empirically adequate, ones whose models fit the observable phenomena, while recognizing that what counts as an observable phenomenon is a function of what the epistemic community is (that *observable* is *observable-to-us*).

The notion of empirical adequacy in this answer will have to be spelt out very carefully if it is not to bite the dust among hackneyed objections. I shall try to do so in the next chapter. But the point stands: even if observability has nothing to do with existence (is, indeed, too anthropocentric for that), it may still have much to do with the proper epistemic attitude to science.

§3. *Inference to the Best Explanation*

A view advanced in different ways by Wilfrid Sellars, J. J. C. Smart, and Gilbert Harman is that the canons of rational inference require scientific realism. If we are to follow the same patterns of inference with respect to this issue as we do in science itself, we shall find ourselves irrational unless we assert the truth of the scientific theories we accept. Thus Sellars says: 'As I see it, to have good reason for holding a theory is *ipso facto* to have good reason for holding that the entities postulated by the theory exist.'[7]

The main rule of inference invoked in arguments of this sort is the rule of *inference to the best explanation*. The idea is perhaps to be credited to C. S. Peirce,[8] but the main recent attempts to explain this rule and its uses have been made by Gilbert Harman.[9] I shall only present a simplified version. Let us suppose that we have evidence E, and are considering several hypotheses, say H and H'. The rule then says that we should infer H rather than H' exactly if H is a better explanation of E than H' is. (Various qualifications are necessary to avoid inconsistency: we should always try to move to the best over-all explanation of all available evidence.)

It is argued that we follow this rule in all 'ordinary' cases; and that if we follow it consistently everywhere, we shall be led to scientific realism, in the way Sellars's dictum suggests. And surely there are many telling 'ordinary' cases: I hear scratching in the wall, the patter of little feet at midnight, my cheese disappears—and I

infer that a mouse has come to live with me. Not merely that these apparent signs of mousely presence will continue, not merely that all the observable phenomena will be as if there is a mouse; but that there really is a mouse.

Will this pattern of inference also lead us to belief in unobservable entities? Is the scientific realist simply someone who consistently follows the rules of inference that we all follow in more mundane contexts? I have two objections to the idea that this is so.

First of all, what is meant by saying that we all *follow* a certain rule of inference? One meaning might be that we deliberately and consciously 'apply' the rule, like a student doing a logic exercise. That meaning is much too literalistic and restrictive; surely all of mankind follows the rules of logic much of the time, while only a fraction can even formulate them. A second meaning is that we act in accordance with the rules in a sense that does not require conscious deliberation. That is not so easy to make precise, since each logical rule is a rule of permission (*modus ponens* allows you to infer *B* from *A* and (if *A then B*), but does not forbid you to infer (*B or A*) instead). However, we might say that a person behaved in accordance with a set of rules in that sense if every conclusion he drew could be reached from his premisses via those rules. But this meaning is much too loose; in this sense we always behave in accordance with the rule that any conclusion may be inferred from any premiss. So it seems that to be following a rule, I must be willing to believe all conclusions it allows, while definitely unwilling to believe conclusions at variance with the ones it allows—or else, change my willingness to believe the premisses in question.

Therefore the statement that we all follow a certain rule in certain cases, is a *psychological hypothesis* about what we are willing and unwilling to do. It is an empirical hypothesis, to be confronted with data, and with rival hypotheses. Here is a rival hypothesis: we are always willing to believe that the theory which best explains the evidence, is empirically adequate (that all the observable phenomena are as the theory says they are).

In this way I can certainly account for the many instances in which a scientist appears to argue for the acceptance of a theory or hypothesis, on the basis of its explanatory success. (A number of such instances are related by Thagard.[8]) For, remember: I equate the acceptance of a scientific theory with the belief that it is empirically adequate. We have therefore two rival hypotheses concerning these instances of scientific inference, and the one is apt in a realist account, the other in an anti-realist account.

Cases like the mouse in the wainscoting cannot provide telling evidence between those rival hypotheses. For the mouse *is* an observable thing; therefore 'there is a mouse in the wainscoting' and 'All observable phenomena are as if there is a mouse in the wainscoting' are totally equivalent; each implies the other (given what we know about mice).

It will be countered that it is less interesting to know whether people do follow a rule of inference than whether they ought to follow it. Granted; but the premiss that we all follow the rule of inference to the best explanation when it comes to mice and other mundane matters—that premiss is shown wanting. It is not warranted by the evidence, because that evidence is not telling *for* the premiss *as against* the alternative hypothesis I proposed, which is a relevant one in this context.

My second objection is that even if we were to grant the correctness (or worthiness) of the rule of inference to the best explanation, the realist needs some further premiss for his argument. For this rule is only one that dictates a choice when given a set of rival hypotheses. In other words, we need to be committed to belief in one of a range of hypotheses before the rule can be applied. Then, under favourable circumstances, it will tell us which of the hypotheses in that range to choose. The realist asks us to choose between different hypotheses that explain the regularities in certain ways; but his opponent always wishes to choose among hypotheses of the form 'theory T_i is empirically adequate'. So the realist will need his special extra premiss that every universal regularity in nature needs an explanation, before the rule will make realists of us all. And that is just the premiss that distinguishes the realist from his opponents (and which I shall examine in more detail in Sections 4 and 5 below).

The logically minded may think that the extra premiss can be bypassed by logical *léger-de-main*. For suppose the data are that all facts observed so far accord with theory T; then T is one possible explanation of those data. A rival is *not-T* (that T is false). This rival is a very poor explanation of the data. So we *always* have a set of rival hypotheses, and the rule of inference to the best explanation leads us unerringly to the conclusion that T is true. Surely I am committed to the view that T is true or T is false?

This sort of epistemological rope-trick does not work of course. To begin, I am committed to the view that T is true or T is false, but not thereby committed to an inferential move to one of the two! The rule operates only if I have decided not to remain neutral between these two possibilities.

Secondly, it is not at all likely that the rule will be applicable to such logically concocted rivals. Harman lists various criteria to apply to the evaluation of hypotheses *qua* explanations.[10] Some are rather vague, like simplicity (but is simplicity not a reason to use a theory whether you believe it or not?). The precise ones come from statistical theory which has lately proved of wonderful use to epistemology:

> H is a better explanation than H' (*ceteris paribus*) of E, provided:
> (a) $P(H) > P(H')$—H has higher probability than H'
> (b) $P(E/H) > P(E/H')$—H bestows higher probability on E than H' does.

The use of 'initial' or *a priori* probabilities in (a)—the initial plausibility of the hypotheses themselves—is typical of the so-called *Bayesians*. More traditional statistical practice suggests only the use of (b). But even that supposes that H and H' bestow definite probabilities on E. If H' is simply the denial of H, that is not generally the case. (Imagine that H says that the probability of E equals $\frac{3}{4}$. The very most that *not-H* will entail is that the probability of E is some number other than $\frac{3}{4}$; and usually it will not even entail that much, since H will have other implications as well.)

Bayesians tend to cut through this 'unavailability of probabilities' problem by hypothesizing that everyone has a specific subjective probability (degree of belief) for every proposition he can formulate. In that case, no matter what E, H, H' are, all these probabilities really are (in principle) available. But they obtain this availability by making the probabilities thoroughly subjective. I do not think that scientific realists wish their conclusions to hinge on the subjectively established initial plausibility of there being unobservable entities, so I doubt that this sort of Bayesian move would help here. (This point will come up again in a more concrete form in connection with an argument by Hilary Putnam.)

I have kept this discussion quite abstract; but more concrete arguments by Sellars, Smart, and Putnam will be examined below. It

should at least be clear that there is no open-and-shut argument from common sense to the unobservable. Merely following the ordinary patterns of inference in science does not obviously and automatically make realists of us all.

§4. *Limits of the Demand for Explanation*

In this section and the next two, I shall examine arguments for realism that point to explanatory power as a criterion for theory choice. That this is indeed a criterion I do not deny. But these arguments for realism succeed only if the demand for explanation is supreme—if the task of science is unfinished, *ipso facto*, as long as any pervasive regularity is left unexplained. I shall object to this line of argument, as found in the writings of Smart, Reichenbach, Salmon, and Sellars, by arguing that such an unlimited demand for explanation leads to a demand for hidden variables, which runs contrary to at least one major school of thought in twentieth-century physics. I do not think that even these philosophers themselves wish to saddle realism with logical links to such consequences: but realist yearnings were born among the mistaken ideals of traditional metaphysics.

In his book *Between Science and Philosophy*, Smart gives two main arguments for realism. One is that only realism can respect the important distinction between *correct* and *merely useful* theories. He calls 'instrumentalist' any view that locates the importance of theories in their use, which requires only empirical adequacy, and not truth. But how can the instrumentalist explain the usefulness of his theories?

> Consider a man (in the sixteenth century) who is a realist about the Copernican hypothesis but instrumentalist about the Ptolemaic one. He can explain the instrumental usefulness of the Ptolemaic system of epicycles because he can prove that the Ptolemaic system can produce almost the same predictions about the apparent motions of the planets as does the Copernican hypothesis. Hence the assumption of the realist truth of the Copernican hypothesis explains the instrumental usefulness of the Ptolemaic one. Such an explanation of the instrumental usefulness of certain theories would not be possible if *all* theories were regarded as merely instrumental.[11]

What exactly is meant by 'such an explanation' in the last sentence? If no theory is assumed to be true, then no theory has its usefulness explained as following from the truth of another one—granted. But would we have less of an explanation of the usefulness of the

Ptolemaic hypothesis if we began instead with the premiss that the Copernican gives implicitly a very accurate description of the motions of the planets as observed from earth? This would not assume the truth of Copernicus's heliocentric hypothesis, but would still entail that Ptolemy's simpler description was also a close approximation of those motions.

However, Smart would no doubt retort that such a response pushes the question only one step back: what explains the accuracy of predictions based on Copernicus's theory? If I say, the empirical adequacy of that theory, I have merely given a verbal explanation. For of course Smart does not mean to limit his question to actual predictions—it really concerns all actual and possible predictions and retrodictions. To put it quite concretely: what explains the fact that all observable planetary phenomena fit Copernicus's theory (if they do)? From the medieval debates, we recall the nominalist response that the basic regularities are merely brute regularities, and have no explanation. So here the anti-realist must similarly say: that the observable phenomena exhibit these regularities, because of which they fit the theory, is merely a brute fact, and may or may not have an explanation in terms of unobservable facts 'behind the phenomena'—it really does not matter to the goodness of the theory, nor to our understanding of the world.

Smart's main line of argument is addressed to exactly this point. In the same chapter he argues as follows. Suppose that we have a theory T which postulates micro-structure directly, and macro-structure indirectly. The statistical and approximate laws about macroscopic phenomena are only partially spelt out perhaps, and in any case derive from the precise (deterministic or statistical) laws about the basic entities. We now consider theory T', which is part of T, and says only what T says about the macroscopic phenomena. (How T' should be characterized I shall leave open, for that does not affect the argument here.) Then he continues:

> I would suggest that the realist could (say) ... that the success of T′ is explained by the fact that the original theory T is true of the things that it is ostensibly about; in other words by the fact that there really are electrons or whatever is postulated by the theory T. If there were no such things, and if T were not true in a realist way, would not the success of T′ be quite inexplicable? One would have to suppose that there were innumerable lucky accidents about the behaviour mentioned in the observational vocabulary, so that they behaved miraculously *as if* they were brought about by the non-existent things ostensibly talked about in the theoretical vocabulary.[12]

In other passages, Smart speaks similarly of 'cosmic coincidences'. The regularities in the observable phenomena must be explained in terms of deeper structure, for otherwise we are left with a belief in lucky accidents and coincidences on a cosmic scale.

I submit that if the demand for explanation implicit in these passages were precisely formulated, it would at once lead to absurdity. For if the mere fact of postulating regularities, without explanation, makes T' a poor theory, T will do no better. If, on the other hand, there is some precise limitation on what sorts of regularities can be postulated as basic, the context of the argument provides no reason to think that T' must automatically fare worse than T.

In any case, it seems to me that it is illegitimate to equate being a lucky accident, or a coincidence, with having no explanation. It was by coincidence that I met my friend in the market—but I can explain why I was there, and he can explain why he came, so together we can explain how this meeting happened. We call it a coincidence, not because the occurrence was inexplicable, but because we did not severally go to the market in order to meet.[13] There cannot be a requirement upon science to provide a theoretical elimination of coincidences, or accidental correlations in general, for that does not even make sense. There is nothing here to motivate the demand for explanation, only a restatement in persuasive terms.

§5. *The Principle of the Common Cause*

Arguing against Smart, I said that if the demand for explanation implicit in his arguments were precisely formulated, it would lead to absurdity. I shall now look at a precise formulation of the demand for explanation: Reichenbach's principle of the common cause. As Salmon has recently pointed out, if this principle is imposed as a demand on our account of what there is in the world, then we are led to postulate the existence of unobservable events and processes.[14]

I will first state the argument, and Reichenbach's principle, in a rough, intuitive form, and then look at its precise formulation. Suppose that two sorts of events are found to have a correlation. A simple example would be that one occurs whenever the other does; but the correlation may only be statistical. There is apparently a significant correlation between heavy cigarette-smoking and cancer, though merely a statistical one. Explaining such a correlation

requires finding what Reichenbach called a *common cause*. But, the argument runs, there are often among observable events no common causes of given observable correlations. Therefore, scientific explanation often requires that there be certain unobservable events.

Reichenbach held it to be a principle of scientific methodology that every statistical correlation (at least, every positive dependence) must be explained through common causes. This means then that the very project of science will necessarily lead to the introduction of unobservable structure behind the phenomena. Scientific explanation will be impossible unless there are unobservable entities; but the aim of science is to provide scientific explanation; therefore, the aim of science can only be served if it is true that there are unobservable entities.

To examine this argument, we must first see how Reichenbach arrived at his notion of common causes and how he made it precise. I will then argue that his principle cannot be a general principle of science at all, and secondly, that the postulation of common causes (when it does occur) is also quite intelligible without scientific realism.

Reichenbach was one of the first philosophers to recognize the radical 'probabilistic turn' of modern physics. The classical ideal of science had been to find a method of description of the world so fine that it could yield deterministic laws for all processes. This means that, if such a description be given of the state of the world (or, more concretely, of a single isolated system) at time t, then its state at later time $t+d$ is uniquely determined. What Reichenbach argued very early on is that this ideal has a factual presupposition: it is not logically necessary that such a fine method of description exists, even in principle.[15] This view became generally accepted with the development of quantum mechanics.

So Reichenbach urged philosophers to abandon that classical ideal as the standard of completeness for a scientific theory. Yet it is clear that, if science does not seek for deterministic laws relating events to what happened before them, it does seek for *some* laws. And so Reichenbach proposed that the correct way to view science is as seeking for 'common causes' of a probabilistic or statistical sort.

We can make this precise using the language of probability theory. Let A and B be two events; we use P to designate their probability of occurrence. Thus $P(A)$ is the probability that A occurs and $P(A\&B)$ the probability that both A and B occur. In addition, we

must consider the probability that A occurs *given that* B occurs. Clearly the probability of rain *given that* the sky is overcast, is higher than the probability of rain in general. We say that B is statistically relevant to A if the probability of A *given* B—written $P(A/B)$—is different from $P(A)$. If $P(A/B)$ is higher than $P(A)$, we say that there is a positive correlation. Provided A and B are events which have some positive likelihood of occurrence (i.e. $P(A)$, $P(B)$ are not zero), this is a symmetric relationship. The precise definitions are these:

(a) the probability of A given B is defined provided $P(B) \neq O$, and is

$$P(A/B) = \frac{P(A\&B)}{P(B)}$$

(b) B is statistically relevant to A exactly if $P(A/B) \neq P(A)$

(c) there is a positive correlation between A and B exactly if $P(A\&B) > P(A) \,.\, P(B)$

(d) from (a) and (c) it follows that, if $P(A) \neq O$ and $P(B) \neq O$, then there is a positive correlation between A and B exactly if

$$P(A/B) > P(A),$$

and also if and only if

$$P(B/A) > P(B)$$

To say therefore that there is a positive correlation between cancer and heavy cigarette-smoking, is to say that the incidence of cancer among heavy cigarette-smokers is greater than it is in the general population. But because of the symmetry of A and B in (d), this statement by itself gives no reason to think that the smoking produces the cancer rather than the cancer producing the smoking, or both being produced by some other factor, or by several other factors, if any.

We are speaking here of facts relating to the same time. The cause we seek in the past: heavy smoking at one time is followed (with certain probabilities) by heavy smoking at a later time, and also by being cancerous at that later time. We have in this past event C really found the *common cause* of this present correlation if

$$P(A/B\&C) = P(A/C)$$

We may put this as follows: relative to the information that C has occurred, A and B are statistically independent. We can define the

probability of an event X, whether by itself or conditional on another event Y, *relative to* C as follows:

(e) the probability relative to C is defined as

$$P_c(X) = P(X/C)$$
$$P_c(X/Y) = P_c(X\&Y) \div P_c(Y)$$
$$= P(X/Y\&C)$$
provided $P_c(Y) \neq O$, $P(C) \neq O$

So to say that C is the common cause for the correlation between A and B is to say that, relative to C there is no such correlation. C explains the correlation, because we notice a correlation only as long as we do not take C into account.

Reichenbach's *Principle of the Common Cause* is that *every* relation of positive statistical relevance must be explained by statistical past common causes, in the above way.[16] To put it quite precisely and in Reichenbach's own terms:

> If coincidences of two events A and B occur more frequently than would correspond to their independent occurrence, that is, if the events satisfy the relation
>
> (1) $P(A\&B) > P(A).P(B)$,
>
> then there exists a common cause C for these events such that the fork ACB is *conjunctive*, that is, satisfies relations (2)–(5) below:
>
> (2) $P(A\&B/C) = P(A/C).P(B/C)$
> (3) $P(A\&B/\bar{C}) = P(A/\bar{C}).P(B/\bar{C})$
> (4) $P(A/C) > P(A/\bar{C})$
> (5) $P(B/C) > P(B/\bar{C})$

(1) follows logically from (2)–(5).

This principle of the common cause is at once precise and persuasive. It may be regarded as a formulation of the conviction that lies behind such arguments as that of Smart, requiring the elimination of 'cosmic coincidence' by science. But it is not a principle that guides twentieth-century science, because it is too close to the demand for deterministic theories of the world that Reichenbach wanted to reject. I shall show this by means of a schematic example; but this example will incorporate the sorts of non-classical correlations which distinguish quantum mechanics from classical physics. I refer here to the correlations exhibited by the thought

experiment of Einstein, Podolski, and Rosen in their famous paper 'Can Quantum-Mechanical Description of Reality be Considered Complete?' These correlations are not merely theoretical: they are found in many actual experiments, such as Compton scattering and photon pair production. I maintain in addition that correlations sufficiently similar to refute the principle of the common cause must appear in almost any indeterministic theory of sufficient complexity.[17] Imagine that you have studied the behaviour of a system or object which, after being in state S, always goes into another state which may be characterized by various attributes $F_1, \ldots, F_n$ and $G_1, \ldots, G_n$. Suppose that you have come to the conclusion that this transition is genuinely indeterministic, but you can propose a theory about the transition probabilities:

(8) (a) $P(F_i/S) = 1/n$ (b) $P(G_i/S) = 1/n$
(c) $P(F_i \equiv G_i/S) = 1$

where $\equiv$ means *if and only if* or *when and exactly when*. In other words, it is pure chance whether the state to which S transits is characterized by a given one of the F-attributes, and similarly for the G-attributes, but certain that it is characterized by F_1 if it is characterized by G_1, by F_2 if by G_2, and so on.

If we are convinced that this is an irreducible, indeterministic phenomenon, so that S is a complete description of the initial state, then we have a violation of the principle of the common cause. For from (8) we can deduce

(9) $P(F_i/S).P(G_i/S) = 1/n^2$
$P(F_i\&G_i/S) = P(F_i/S) = 1/n$

which numbers are equal only if n is zero or one—the deterministic case. In all other cases, S does not qualify as the common cause of the new state's being F_i and G_i, and if S is complete, nothing else can qualify either.

The example I have given is schematic and simplified, and besides its indeterminism, it also exhibits a certain discontinuity, in that we discuss the transition of a system from one state S into a new state. In classical physics, if a physical quantity changed its value from i to j it would do so by taking on all the values between i and j in succession, that is, changing continuously. Would Reichenbach's principle be obeyed at least in some non-trivial, indeterministic theory in which all quantities have a continuous spectrum of values

and all change is continuous? I think not, but I shall not argue this further. The question is really academic, for if the principle requires that, then it is also not acceptable to modern physical science.

Could one change a theory which violates Reichenbach's principle into one that obeys it, without upsetting its empirical adequacy? Possibly; one would have to deny that the attribution of state S gives complete information about the system at the time in question, and postulate *hidden parameters* that underlie these states. Attempts to do so for quantum mechanics are referred to as *hidden variable theories*, but it can be shown that if such a theory is empirically equivalent to orthodox quantum mechanics, then it still exhibits non-local correlations of a non-classical sort, which would still violate Reichenbach's principle. But again, the question is academic, since modern physics does not recognize the need for such hidden variables.

Could Reichenbach's principle be weakened so as to preserve its motivating spirit, while eliminating its present unacceptable consequences? As part of a larger theory of explanation (which I shall discuss later), Wesley Salmon has proposed to disjoin equation (2) above with

$$(2^*) \qquad P(A\&B/C) > P(A/C) \cdot P(B/C)$$

in which case C would still qualify as common cause. Note that in the schematic example I gave, S would then qualify as a common cause for the events F_i and G_i.

But so formulated, the principle yields a regress. For suppose (2*) is true. Then we note a positive correlation *relative to* C:

$$P_c(A\&B) > P_c(A) \cdot P_c(B)$$

to which the principle applies and for which it demands a common cause C'. This regress stops only if, at some point, the exhibited common cause satisfies the original equation (2), which brings us back to our original situation; or if some other principle is used to curtail the demand for explanation.

In any case, weakening the principle in various ways (and certainly it will have to be weakened if it is going to be acceptable in any sense) will remove the force of the realist arguments. For any weakening is an agreement to leave some sorts of 'cosmic coincidence' unexplained. But that is to admit the tenability of the nominalist/empiricist point of view, for the demand for explanation ceases then to be a scientific 'categorical imperative'.

Nevertheless, there is a problem here that should be faced. Without a doubt, many scientific enterprises can be characterized as searches for common causes to explain correlations. What is the anti-realist to make of this? Are they not searches for explanatory realities behind the phenomena?

I think that there are two senses in which a principle of common causes is operative in the scientific enterprise, and both are perfectly intelligible without realism.

To the anti-realist, all scientific activity is ultimately aimed at greater knowledge of what is observable. So he can make sense of a search for common causes only if that search aids the acquisition of that sort of knowledge. But surely it does! When past heavy smoking is postulated as a causal factor for cancer, this suggests a further correlation between cancer and either irritation of the lungs, or the presence of such chemicals as nicotine in the bloodstream, or both. The postulate will be vindicated if such suggested further correlations are indeed found, and will, if so, have aided in the search for larger scale correlations among observable events.[18] This view reduces the Principle of Common Cause from a regulative principle for all scientific activity to one of its tactical maxims.

There is a second sense in which the principle of the common cause may be operative: as advice for the construction of theories and models. One way to construct a model for a set of observable correlations is to exhibit hidden variables with which the observed ones are individually correlated. This is a theoretical enterprise, requiring mathematical embedding or existence proofs. But if the resulting theory is then claimed to be empirically adequate, there is no claim that all aspects of the model correspond to 'elements of reality'. As a theoretical directive, or as a practical maxim, the principle of the common cause may well be operative in science—but not as a demand for explanation which would produce the metaphysical baggage of hidden parameters that carry no new empirical import.

§6. *Limits to Explanation: a Thought Experiment*

Wilfrid Sellars was one of the leaders of the return to realism in

philosophy of science and has, in his writings of the past three decades, developed a systematic and coherent scientific realism. I have discussed a number of his views and arguments elsewhere; but will here concentrate on some aspects that are closely related to the arguments of Smart, Reichenbach, and Salmon just examined.[19] Let me begin by setting the stage in the way Sellars does.

There is a certain over-simplified picture of science, the 'levels picture', which pervades positivist writings and which Sellars successfully demolished.[20] In that picture, singular observable facts ('this crow is black') are scientifically explained by general observable regularities ('all crows are black') which in turn are explained by highly theoretical hypotheses not restricted in what they say to the observable. The three levels are commonly called those of *fact*, of *empirical law*, and of *theory*. But, as Sellars points out, theories do not explain, or even entail such empirical laws—they only show why observable things obey these so-called laws to the extent they do.[21] Indeed, perhaps we have no such empirical laws at all: all crows are black—except albinos; water boils at 100°C—provided atmospheric pressure is normal; a falling body accelerates—provided it is not intercepted, or attached to an aeroplane by a static line; and so forth. On the level of the observable we are liable to find only putative laws heavily subject to unwritten *ceteris paribus* qualifications.

This is, so far, only a methodological point. We do not really expect theories to 'save' our common everyday generalizations, for we ourselves have no confidence in their strict universality. But a theory which says that the micro-structure of things is subject to *some* exact, universal regularities, must imply the same for those things themselves. This, at least, is my reaction to the points so far. Sellars, however, sees an inherent inferiority in the description of the observable alone, an incompleteness which requires (*sub specie* the aims of science) an introduction of an unobservable reality behind the phenomena. This is brought out by an interesting 'thought-experiment'.

Imagine that at some early stage of chemistry it had been found that different samples of gold dissolve in *aqua regia* at different rates, although 'as far as can be observationally determined, the specimens and circumstances are identical'.[22] Imagine further that the response of chemistry to this problem was to postulate two distinct micro-

structures for the different samples of gold. Observationally unpredictable variation in the rate of dissolution is explained by saying that the samples are mixtures (not compounds) of these two (observationally identical) substances, each of which has a fixed rate of dissolution.

In this case we have explanation through laws which have no observational counterparts that can play the same role. Indeed, no explanation seems possible unless we agree to find our physical variables outside the observable. But science aims to explain, must try to explain, and so must require a belief in this unobservable micro-structure. So Sellars contends.

There are at least three questions before us. Did this postulation of micro-structure really have no new consequences for the observable phenomena? Is there really such a demand upon science that it must explain—even if the means of explanation bring no gain in empirical predictions? And thirdly, could a *different* rationale exist for the use of a micro-structure picture in the development of a scientific theory in a case like this?

First, it seems to me that these hypothetical chemists did postulate new observable regularities as well. Suppose the two substances are A and B, with dissolving rates x and $x+y$ and that every gold sample is a mixture of these substances. Then it follows that every gold sample dissolves at a rate no lower than x and no higher than $x+y$; *and* that between these two any value may be found—to within the limits of accuracy of gold mixing. None of this is implied by the data that different samples of gold have dissolved at various rates between x and $x+y$. So Sellar's first contention is false.

We may assume, for the sake of Sellars's example, that there is still no way of predicting dissolving rates any further. Is there then a categorical demand upon science to explain this variation which does not depend on other observable factors? We have seen that a precise version of such a demand (Reichenbach's principle of the common cause) could result automatically in a demand for hidden variables, providing a 'classical' underpinning for indeterministic theories. Sellars recognized very well that a demand for hidden variables would run counter to the main opinions current in quantum physics. Accordingly he mentions '.... the familiar point that the irreducibly and lawfully statistical ensembles of quantum-mechanical theory are mathematically inconsistent with the

assumption of hidden variables.'[23] Thus, he restricts the demand for explanation, in effect, to just those cases where it is *consistent* to add hidden variables to the theory. And consistency is surely a logical stopping-point.

This restriction unfortunately does not prevent the disaster. For while there are a number of proofs that hidden variables cannot be supplied so as to turn quantum mechanics into a classical sort of deterministic theory, those proofs are based on requirements much stronger than consistency. To give an example, one such assumption is that two distinct physical variables cannot have the same statistical distributions in measurement on all possible states.[24] Thus it is assumed that, if we cannot point to some possible difference in empirical predictions, then there is no real difference at all. If such requirements were lifted, and consistency alone were the criterion, hidden variables could indeed be introduced. I think we must conclude that science, in contrast to scientific realism, does not place an overriding value on explanation in the absence of any gain for empirical results.

Thirdly, then, let us consider how an anti-realist could make sense of those hypothetical chemists' procedure. After pointing to the new empirical implications which I mentioned three paragraphs ago, he would point to methodological reasons. By imagining a certain sort of micro-structure for gold and other metals, say, we might arrive at a theory governing many observationally disparate substances; and this might then have implications for new, wider empirical regularities when such substances interact. This would only be a hope, of course; no hypothesis is guaranteed to be fruitful—but the point is that the true demand on science is not for explanation *as such*, but for imaginative pictures which have a hope of suggesting new statements of observable regularities and of correcting old ones. This point is exactly the same as that for the principle of the common cause.

§7. *Demons and the Ultimate Argument*

Hilary Putnam, in the course of his discussions of realism in logic and mathematics, advanced several arguments for scientific realism as well. In *Philosophy of Logic* he concentrates largely on indispensability arguments—concepts of mathematical entities are indispensable to non-elementary mathematics, theoretical concepts are indispensable to physics.[25] Then he confronts the philosophical

position of Fictionalism, which he gleans from the writings of Vaihinger and Duhem:

> (T)he fictionalist says, in substance, 'Yes, certain concepts ... are indispensable, but no, that has no tendency to show that entities corresponding to those concepts actually exist. It only shows that those 'entities' are useful *fictions*'.[26]

Glossed in terms of theories: even if certain kinds of theories are indispensable for the advance of science, that does not show that those theories are true *in toto*, as well as empirically correct.

Putnam attacks this position in a roundabout way, first criticizing bad arguments against Fictionalism, and then garnering his reasons for rejecting Fictionalism from that discussion. The main bad reason he sees is that of Verificationism. The logical positivists adhered to the verificationist theory of meaning; which is roughly that the total cognitive content of an assertion, all that is meaningful in it, is a function of what empirical results would verify or refute it. Hence, they would say that there are no real differences between two hypotheses with the same empirical content. Consider two theories of what the world is like: Rutherford's atomic theory, and Vaihinger's hypothesis that, although perhaps there are no electrons and such, the observable world is nevertheless exactly as if Rutherford's theory were true. The Verificationist would say: these two theories, although Vaihinger's appears to be consistent with the denial of Rutherford's, amount to exactly the same thing.

Well, they don't, because the one says that there are electrons, and the other allows that there may not be. Even if the observable phenomena are as Rutherford says, the unobservable may be different. However, the positivists would say, if you argue that way, then you will automatically become a prey to scepticism. You will have to admit that there are possibilities you cannot prove or disprove by experiment, and so you will have to say that we just cannot know what the world is like. Worse; you will have no reason to reject any number of outlandish possibilities; demons, witchcraft, hidden powers collaborating to fantastic ends.

Putnam considers this argument for Verificationism to be mistaken, and his answer to it, strangely enough, will also yield an answer to the Fictionalism rejected by the verificationist. To dispel the bogey of scepticism, Putnam gives us a capsule introduction to contemporary (Bayesian) epistemology: Rationality requires that if

two hypotheses have all the same testable consequences (consequences for evidence that could be gathered), then we should not accept the one which is *a priori the less plausible*. Where do we get our *a priori* plausibility orderings? These we supply ourselves, either individually or as communities: to accept a plausibility ordering is neither

> to make a judgment of empirical fact nor to state a theorem of deductive logic; it is to take a methodological stand. One can only say whether the demon hypothesis is 'crazy' or not if one has taken such a stand; I report the stand I have taken (and, speaking as one who has taken this stand, I add: and the stand all rational men take, implicitly or explicitly).[27]

On this view, the difference between Rutherford and Vaihinger, or between Putnam and Duhem, is that (although they presumably agree on the implausibility of demons) they disagree on the *a priori* plausibility of electrons. Does each simply report the stand he has taken, and add: this is, in my view, the stand of all rational men? How disappointing.

Actually, it does not quite go that way. Putnam has skilfully switched the discussion from electrons to demons, and asked us to consider how we could rule out their existence. As presented, however, Vaihinger's view differed from Rutherford's by being logically weaker—it only withheld assent to an existence assertion. It follows automatically that Vaihinger's view cannot be *a priori* less plausible than Rutherford's. Putnam's ideological manœuvre could at most be used to accuse an 'atheistic' anti-realist of irrationality (relative to Putnam's own stand, of course)—not one of the agnostic variety.

Putnam concludes this line of reasoning by asking what more could be wanted as evidence for the truth of a theory than what the realist considers sufficient: 'But then . . . what *further* reasons could one want before one regarded it as rational to *believe* a theory?'[28] The answer is: *none*—at least if he equates reasons here either with empirical evidence or with compelling arguments. (Inclining reasons are perhaps another matter, especially because Putnam uses the phrase 'rational to believe' rather than 'irrational not to believe'.) Since Putnam has just done us the service of refuting Verificationism, this answer 'none' cannot convict us of irrationality. He has himself just argued forcefully that theories could agree in empirical content and differ in truth-value. Hence, a realist will have

to make a leap of faith. The decision to leap is subject to rational scrutiny, but not *dictated* by reason and evidence.

In a further paper, 'What is Mathematical Truth', Putnam continues the discussion of scientific realism, and gives what I shall call the *Ultimate Argument*. He begins with a formulation of realism which he says he learned from Michael Dummett:

> A realist (with respect to a given theory or discourse) holds that (1) the sentences of that theory are true or false; and (2) that what makes them true or false is something external—that is to say, it is not (in general) our sense data, actual or potential, or the structure of our minds, or our language, etc.[29]

This formulation is quite different from the one I have given even if we instantiate it to the case in which that theory or discourse is science or scientific discourse. Because the wide discussion of Dummett's views has given some currency to his usage of these terms, and because Putnam begins his discussion in this way, we need to look carefully at this formulation.

In my view, Dummett's usage is quite idiosyncratic. Putnam's statement, though very brief, is essentially accurate. In his 'Realism', Dummett begins by describing various sorts of realism in the traditional fashion, as disputes over whether there really exist entities of a particular type. But he says that in some cases he wishes to discuss, such as the reality of the past and intuitionism in mathematics, the central issues seem to him to be about other questions. For this reason he proposes a new usage: he will take such disputes

> as relating, not to a class of entities or a class of terms, but to a class of *statements* ... Realism I characterize as the belief that statements of the disputed class possess an objective truth-value, independently of our means of knowing it: they are true or false in virtue of a reality existing independently of us. The anti-realist opposes to this the view that statements of the disputed class are to be understood only by reference to the sort of thing which we count as evidence for a statement of that class.[30]

Dummett himself notes at once that nominalists are realists in this sense.[31] If, for example, you say that abstract entities do not exist, and sets are abstract entities, hence sets do not exist, then you will certainly accord a truth-value to every statement of set theory. It might be objected that if you take this position then you have a decision procedure for determining the truth-values of these statements (*false* for existentially quantified ones, *true* for universal ones, apply truth tables for the rest). Does that not mean that, on your

view, the truth-values are not independent of our knowledge? Not at all; for you clearly believe that if we had not existed, and *a fortiori* had had no knowledge, the state of affairs with respect to abstract entities would be the same.

Has Dummett perhaps only laid down a necessary condition for realism, in his definition, for the sake of generality? I do not think so. In discussions of quantum mechanics we come across the view that the particles of microphysics are real, and obey the principles of the theory, but at any time t when 'particle x has exact momentum p' is true then 'particle x has position q' is neither true nor false. In any traditional sense, this is a realist position with respect to quantum mechanics.

We note also that Dummett has, at least in this passage, taken no care to exclude non-literal construals of the theory, as long as they are truth–valued. The two are not the same; when Strawson construed 'The king of France in 1905 is bald' as neither true nor false, he was not giving a non-literal construal of our language. On the other hand, people tend to fall back on non-literal construals typically in order to be able to say, 'properly construed, the theory is true.'[32]

Perhaps Dummett is right in his assertion that what is really at stake, in realist disputes of various sorts, is questions about language—or, if not really at stake, at least the only serious philosophical problems in those neighbourhoods. Certainly the arguments in which he engages are profound, serious, and worthy of our attention. But it seems to me that his terminology ill accords with the traditional one. Certainly I wish to define scientific realism so that it need not imply that all statements in the theoretical language are true or false (only that they are all capable of being true or false, that is, there are conditions for each under which it has a truth-value); to imply nevertheless that the aim is that the theories should be true. And the contrary position of constructive empiricism is not anti-realist in Dummett's sense, since it also assumes scientific statements to have truth-conditions entirely independent of human activity or knowledge. But then, I do not conceive the dispute as being about language at all.

In any case Putnam himself does not stick with this weak formulation of Dummett's. A little later in the paper he directs himself to scientific realism *per se*, and formulates it in terms borrowed, he says, from Richard Boyd. The new formulation comes in the course of a

new argument for scientific realism, which I shall call the Ultimate Argument:

> the positive argument for realism is that it is the only philosophy that doesn't make the success of science a miracle. That terms in mature scientific theories typically refer (this formulation is due to Richard Boyd), that the theories accepted in a mature science are typically approximately true, that the same term can refer to the same thing even when it occurs in different theories—these statements are viewed by the scientific realist not as necessary truths but as part of the only scientific explanation of the success of science, and hence as part of any adequate scientific description of science and its relations to its objects.[33]

Science, apparently, is required to explain its own success. There is this regularity in the world, that scientific predictions are regularly fulfilled; and this regularity, too, needs an explanation. Once *that* is supplied we may perhaps hope to have reached the *terminus de jure*?

The explanation provided is a very traditional one—*adequatio ad rem*, the 'adequacy' of the theory to its objects, a kind of mirroring of the structure of things by the structure of ideas—Aquinas would have felt quite at home with it.

Well, let us accept for now this demand for a scientific explanation of the success of science. Let us also resist construing it as merely a restatement of Smart's 'cosmic coincidence' argument, and view it instead as the question why we have successful scientific theories at all. Will this realist explanation with the Scholastic look be a scientifically acceptable answer? I would like to point out that science is a biological phenomenon, an activity by one kind of organism which facilitates its interaction with the environment. And this makes me think that a very different kind of scientific explanation is required.

I can best make the point by contrasting two accounts of the mouse who runs from its enemy, the cat. St. Augustine already remarked on this phenomenon, and provided an intentional explanation: the mouse *perceives that* the cat is its enemy, hence the mouse runs. What is postulated here is the 'adequacy' of the mouse's thought to the order of nature: the relation of enmity is correctly reflected in his mind. But the Darwinist says: Do not ask why the *mouse* runs from its enemy. Species which did not cope with their natural enemies no longer exist. That is why there are only ones who do.

In just the same way, I claim that the success of current scientific theories is no miracle. It is not even surprising to the scientific (Darwinist) mind. For any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive—the ones which *in fact* latched on to actual regularities in nature.[34]

3

To Save The Phenomena[1]

> Physicists call a theory satisfactory if (1) it agrees with the experimental facts, (2) it is logically consistent, and (3) it is simple as compared to other explanations ... In fact, the author's interest in hidden-variable theories was kindled only when recently he became aware of the possibility of such experimental tests.
>
> On the other hand, we do not want to ignore the metaphysical implications of the theory.
>
> F. J. Belinfante, Foreword, *A Survey of Hidden-Variable Theories* (1973)

THE realist arguments discussed so far were developed mainly in a critique of logical positivism. Much of that critique was correct and successful: the positivist picture of science no longer seems tenable. Since that was essentially the only picture of science within philosophical ken, it is imperative to develop a new account of the structure of science. This account should especially provide a new answer to the question: what is the *empirical content* of a scientific theory?

§1. *Models*

Before turning to examples, let us distinguish the syntactic approach to theories from the semantic one which I favour. Modern axiomatics stems from the discussion of alternative geometric theories, which followed the development of non-Euclidean geometry in the nineteenth century. The first meta-mathematics was meta-geometry (a term already used in Bertrand Russell's *Essays on the Foundations of Geometry* in 1897). It will be easiest perhaps to introduce the relevant axiomatic concepts by way of some simple geometric theories. Consider the axioms:[2]

*A*0 There is at least one line.